UNIVERSITÉ DE MONTRÉAL

LOCOMOTIVE MAINTENANCE SCHEDULING AND TRAIN ASSIGNMENT *

by

Jacques Desrosiers and François Soumis

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

LOCOMOTIVE MAINTENANCE SCHEDULING AND TRAIN ASSIGNMENT *

by

Jacques Desrosiers and François Soumis

* Cette étude a été publiée en français, au Centre de recherche sur les transports sous le même numéro de publication et à l'Ecole des Hautes Etudes Commerciales : rapport de recherche #82-04. Titre de la version française : MAINTENANCE DES LOCOMOTIVES.

Ce rapport a été publié grâce à une subvention du fonds F.C.A.C. pour l'aide et le soutien à la recherche.

Université de Montréal

Centre de recherche sur les transports - Publication #239
Ecole des Hautes Etudes Commerciales - Rapport de recherche #82-04
(pour la version française seulement)

Dépôt légal - 3e trimestre 1982

SOMMAIRE

L'exploitation d'un réseau ferroviaire pose le problème de choisir la locomotive pour remorquer chaque train prévu à l'horaire de façon à ce que les locomotives retournent périodiquement à l'atelier de réparation. Nous proposons de traiter ce problème par courte période de planification; cette approche permet de s'adapter aux perturbations et de poser des modèles mathématiques propres à la résolution. Le premier modèle inclue les limites de kilométrage sous forme de contraintes; la solution est obtenue soit par un branch-and -bound où l'on résout successivement des problèmes de flots sur un réseau espace-temps, soit par une décomposition de Dantzig-Wolfe (problème de recouvrement et chemin le plus court avec contraintes de kilométrage sur les noeuds). Le second modèle ajoute à l'objectif des termes de pénalités; pour résoudre, il faut minimiser un objectif non-linéaire et différentiable par morceau, sujet à des contraintes de réseau.

Mots clés: Scheduling, réseau espace-temps, contraintes de distance sur les noeuds.

ABSTRACT

In the operation of a rail network, the problem of assigning locomotives to trains in the schedule must take into account the fact that locomotives must be periodically brought back to the maintenance workshop for repairs. We propose dealing with this problem by using short planning periods; this approach provides flexibility to deal with disturbances and enables suitable mathematical models to be formulated. The first model deals with limits on the distance travelled between maintenance checks by including them as constraints; the solution is obtained either by branch-and-bound where flow problems are successively solved on a space-time network, or by Dantzig-Wolfe decomposition (covering problem and shortest path with distance constraints on the nodes). The second model includes penalty terms in the objective function; to solve this, a non-linear piece-wise differentiable objective function must be minimized subject to network constraints.

Keywords: Scheduling, space-time network, distance interval constraints.

CONTENTS

1. Introduction 1

2. Mathematical Structure of the problem. 2
   2.1 Assignment variables between trains 2
   2.2 Assignment variables in the maintenance workshop. 4
   2.3 Distance variables. 5
   2.4 Constraints 5
   2.5 The objective function. 6

3. Models and solution approach 7
   3.1 Model with distance constraints 7
   3.2 Model with distance terms in the objective function 9

4. Conclusion 12

References 12

1. INTRODUCTION

In the operation of a rail network, the assignment of locomotives to trains in the schedule must take into account the fact that each locomotive must be brought back to its base city periodically so that maintenance work can be performed. At the same time, maintenance and light running costs must be minimized, while excessive variability in the work-load of the maintenance shop must be avoided and the recommended distance travelled between maintenance checks must not be greatly exceeded.

In concrete terms, choices have to be made when several locomotives are located in the same city at the same time. The decision must be made as to how the locomotives will be assigned to trains leaving the city. These assignments determine the paths of the locomotives in the network, the dates of arrival at the base, kilometers travelled and the possibilities of maintenance for each locomotive.

At present, these assignment decisions are left to operating managers who are often reduced to taking short-term decisions to bring a locomotive back to the shop for maintenance as quickly as possible. It is difficult to predict the future impact of a decision. The operating personnel must rely on their own experience in the absence of any tools which would enable them to evaluate the global impact of a decision with respect to the objectives of maintenance regularity and work-load stabilization. Attempts to use operations research to improve the optimality of the solution usually deal with the problem by establishing a period of several weeks in which each locomotive completes a cycle with one or more stops at the maintenance shop. Each cycle is constructed to ensure regular maintenance of each locomotive and the cycles are synchronized to stabilize work-load in the workshop. In spite of its theoretical interest, this approach is difficult to apply in practice because of disturbances due to holidays, equipment breakdowns and changes in train schedules. It is often just as difficult to get back to the predetermined cycles after a disturbance as it is to solve the complete problem.

We propose instead to deal with the problem by using short time periods (probably one week) and to find a new solution for each period by considering the available capacity of the workshop during this period, the accumulated kilometers travelled by each locomotive, and its position in the network at the beginning of the period. For each period, the problem can be solved while taking into account information for a small number of future periods such as planned workshop shut-downs or peak periods requiring maximum circulation of equipment; a horizon of two or three weeks would be chosen. The solution can be obtained using mathematical programming by minimizing the costs for the period while ensuring that a maintenance backlog does not build up. This approach deals with all past and foreseeable disturbances so as to produce a directly applicable optimal solution for this situation. If several new distrubances occur during the week and upset the maintenance schedule, the problem merely has to be solved again to determine how assignments should be made given these distrubances. The short planning period considered reduces problem size and should enable a good solution to be obtained in a reasonable time, even on a small computer.

2. MATHEMATICAL STRUCTURE OF THE PROBLEM

2.1 ASSIGNMENT VARIABLES BETWEEN TRAINS

Let I be the set of arriving trains which free up a locomotive for the period considered. A locomotive can be made available during the period considered or just before and be still available at the beginning of the period. Trains which free up a locomotive at the end of the period for assignment during the following period are not included in I. A train number which is repeated each day is considered as several distinct trains. If an empty train is hauled by another train, these are considered as two distinct trains.

Let J be the set of trains requiring a locomotive during the period considered.

$i \in I$, $j \in J$ are trains.

x_{ij} is a binary variable where $x_{ij} = 1$ when the locomotive from train i is assigned to train j. This variable is defined if the locomotive from train i is available in time to be assigned to train j, and another locomotive is also available.

The assignment variables must be defined when at least two trains are present in the same city at the same time. The figures on the vertical arcs in figure 1 indicating the stock of locomotives in a city show that such situations occur in the example given. Assignment variables must therefore be defined between the train arrivals which contribute to increasing the stock and the departures which reduce it to zero. The zero stock situation in a city separates out assignment problems. It always occurs at least once during the planning period, and usually happens every day, otherwise a locomotive would always be stationary in the city, which would be a non-optimal situation. Between two zero stock situations, the assignment problem between arriving and departing trains is not complete because some trains arrive too late to be assigned to the first departures: here, for example, train 4 cannot be assigned to train 3. Other assignment variables can be eliminated if they are not to be used. If there are preferences for certain assignments, corresponding costs can be attached to the assignments so that preferences are respected.

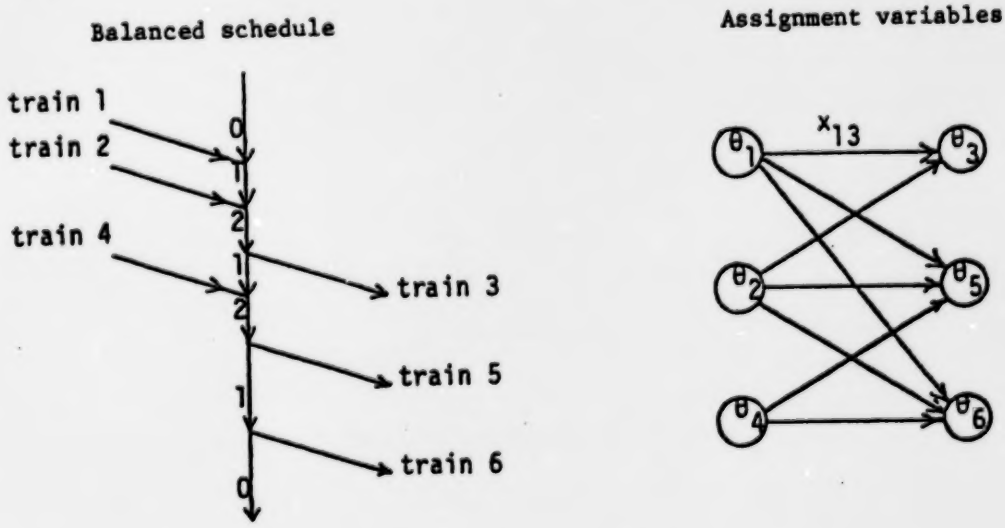


Figure 1: Example of variable definition

2.2 ASSIGNMENT VARIABLES IN THE MAINTENANCE WORKSHOP

At the base city, assignment variables between trains are defined, as well as some new variables y_{ik} and y_{kj}.

K The set of time intervals during which maintenance can be done on a locomotive in the workshop.

$k \in K$ A maintenance time interval.

y_{ik} is a binary variable which takes the value 1 when the locomotive from train i is assigned to the kth maintenance time interval. This variable is defined if train i arrives at the base in time to go to the workshop during interval k.

y_{kj} is a binary variable taking the value 1 when the locomotive leaving the workshop after interval k is assigned to train j. This variable is defined if train j departs from the base after the end of the kth maintenance time interval.

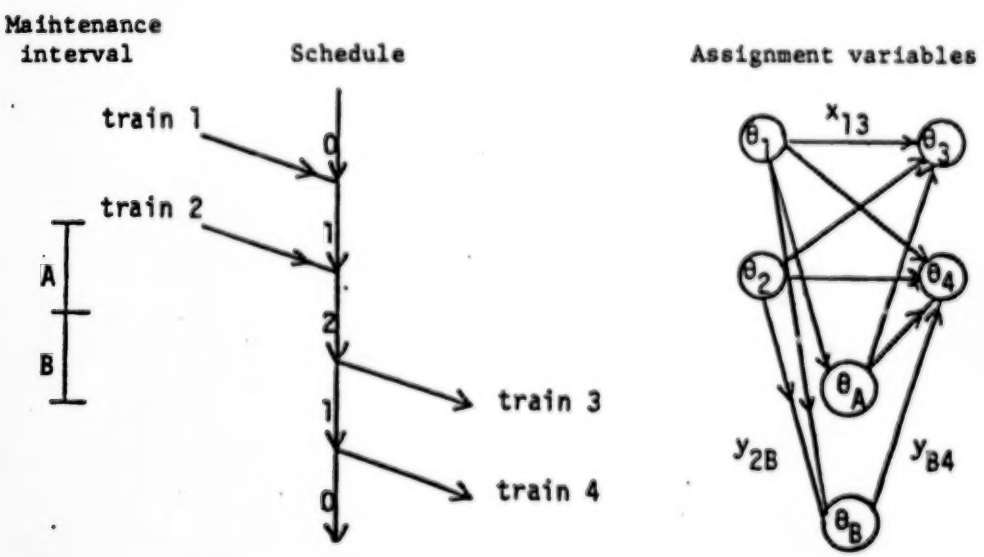


Figure 2: Example of variable definition at base

2.3 DISTANCE VARIABLES

A distance variable is associated with each node in the network.

θ_i, $i \in J$ is a distance variable indicating the kilometers travelled since the last maintenance check by the locomotive made available by train i.

H is the set of trains freeing up a locomotive at the end of the planning period for assignment during the next period.

θ_h, $h \in H$ are distance variables for locomotives made available at the end of the planning period.

θ_k, $k \in K$ are distance variables for locomotives on their arrival at the workshop. These variables therefore indicate the distance travelled between maintenance checks.

d_j, $j \in J$ is the number of kilometers travelled by train j.

2.4 CONSTRAINTS

This notation enables the constraints of the problem to be formulated. First, the assignment constraints are given by:

$$\sum_j x_{ij} = 1 \qquad (\forall i) \qquad (1)$$

$$\sum_i x_{ij} = 1 \qquad (\forall j) \qquad (2)$$

$$\sum_i y_{ik} \leq 1 \qquad (\forall k) \qquad (3)$$

$$\sum_i y_{ik} = \sum_j y_{kj} \qquad (\forall k) \qquad (4)$$

The first constraint indicates that a locomotive made available must be assigned to another train. The second indicates that a locomotive is assigned to each train requiring one for the period considered. The third constraint requires that there be at most one locomotive in the workshop at a time. The fourth corresponds to flow conservation of locomotives in the workshop.

The second group of constraints describes the relationships between distance variables and assignment variables.

$$x_{ij} = 1 \Rightarrow \theta_j = \theta_i + d_j \qquad (\forall ij) \qquad (5)$$

$$y_{ik} = 1 \Rightarrow \theta_k = \theta_i \qquad (\forall ik) \qquad (6)$$

$$y_{kj} = 1 \Rightarrow \theta_j = d_j \qquad (\forall kj) \qquad (7)$$

Constraint (5) indicates that if the locomotive from train i is assigned to train j, the distance travelled by this locomotive after it has hauled train j will be increased by the distance d_j travelled since it was made available by train i. Constraint (6) indicates that if the locomotive made available by train i goes to the workshop during interval k, the distance travelled on entry into the workshop is the same as after train i. Constraint (7) indicates that if the locomotive leaving the workshop at the end of interval k is assigned to train j, the distance θ_j travelled after train j is equal to d_j, as the cumulated distance is reset to zero on leaving the workshop.

2.5 THE OBJECTIVE FUNCTION

The objective is to minimize maintenance costs

$$\text{Min } C * \sum_i \sum_k y_{ik}$$

The constant C is the cost per locomotive and the sum of the y_{ik} gives the number of locomotives which have visited the workshop during the planning period.

The formulation must be completed to ensure that savings are not achieved by unduly spacing out the maintenance checks and putting them off to following periods. This can be avoided by adding constraints or by adding terms to the objective function. The two possibilities are presented here.

3. MODELS AND SOLUTION APPROACH

3.1 MODEL WITH DISTANCE CONSTRAINTS

The requirement to undertake periodic maintenance checks (every 5000 km for example) can be expressed in the form of minimum and maximum cumulated distances between which maintenance must take place. As the θ_k, $k \in K$ represent kilometers accumulated at maintenance, the constraint can be expressed as follows:

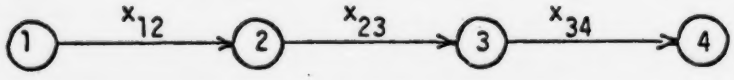
$$M_1 \leq \theta_k \leq M_2 \qquad k \in K \tag{8}$$

In addition, it must be possible to complete work that has been postponed within the distance limits. A maximum can be imposed on the cumulated distance travelled by locomotives ending the period in a given city to ensure that they will be able to return to the base during the next period without exceeding the maximum distance allowed. The maximum for each city will be an inverse function of its distance from the maintenance base:

$$0 \leq \theta_h \leq M_h \qquad h \in H. \tag{9}$$

This model has the same structure as the routing problem with scheduling constraints discussed in Desrosiers (1981a): distance interval constraints take the place of scheduling constraints. Two approaches, using branch-and-bound techniques have been successfully tested for these problems; they apply to this formulation.

The first approach presented in figure 3, solves at each node of the branch-and-bound tree a network flow problem defined by relations 1, 2, 3 and 4 without considering distance constraints. The solutions are integer but infeasible with respect to constraints 5, 6, 7, 8 and 9. Branching is done by introducing or eliminating an arc to achieve feasibility.



distance constraint not satisfied by θ_4

Branching to satisfy distance constraints

R: network without distance constraints (integer solutions)

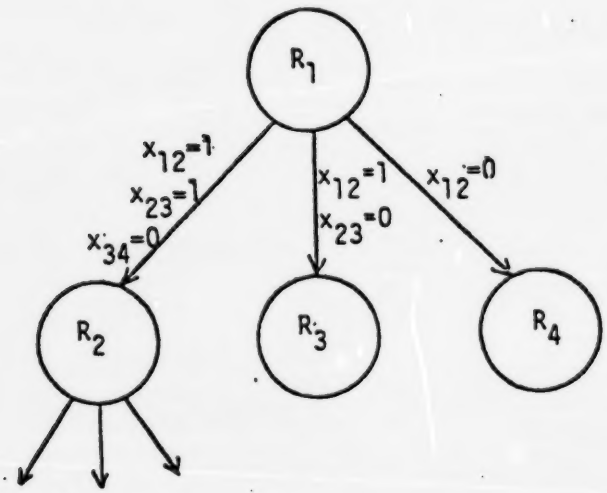
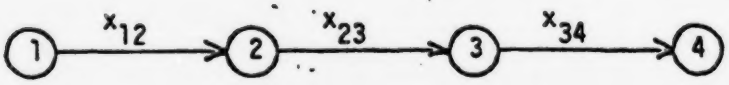


Figure 3: Branch-and-bound tree relaxing distance constraints

This first approach can deal with very large problems as only network problems need be solved. One or more integer solutions are certain to be obtained rapidly. On the other hand, it may be necessary to investigate a very large tree before the optimal solution is obtained. This is the case when several distance constraints are active in a feasible solution: it then becomes necessary to descend low down into the tree to obtain feasibility.

In the second approach presented in figure 4, the whole problem is solved at each node but fractional solutions are obtained for the assignment variables x_{ij}. Branching is based on the addition or elimination of arcs in order to obtain an integer solution.



$0 < x_{12} = x_{23} = x_{34} < 1$; fractional solution

Branching to achieve integrality

RK: network with distance constraints (fractional solutions)

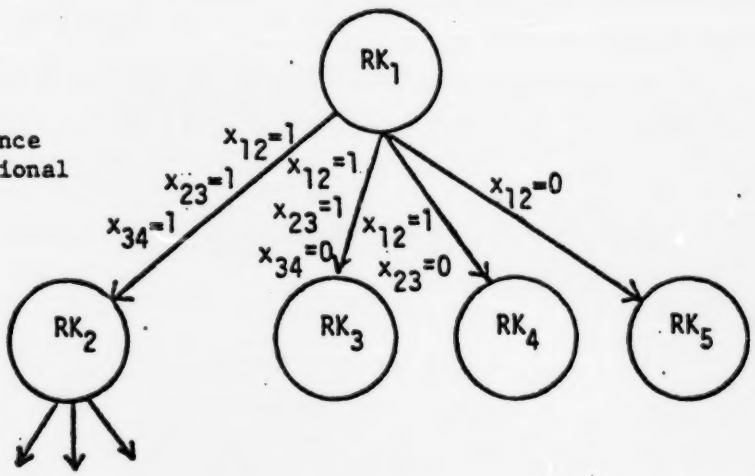


Figure 4: Branch-and-bound tree relaxing integrality constraints

The second approach requires much more work at each node. The network problem with distance constraints is solved by column generation using the simplex algorithm for the covering problem and an adaptation of the shortest path method to construct routes satisfying the distance constraints Desrosiers (1981b). This method is more costly and has large memory requirements for large problems. On the other hand, the tree to be investigated to obtain the integer solution is usually fairly small as a very good approximation of the entire problem is solved at each node.

3.2 MODEL WITH DISTANCE TERMS IN THE OBJECTIVE FUNCTION

The requirement that maintenance be done at regular intervals can also be dealt with by adding penalty terms to the objective function rather than by using constraints (8) and (9). The objective function then becomes:

$$\text{Min } C * \sum_i \sum_k y_{ik} + \sum_{k \in K} f(\theta_k) + g\left(\sum_{h \in H} \theta_h\right) \qquad (10)$$

Recall that the θ_k, $k \in K$ are distances travelled since the last maintenance stop for locomotives arriving at the workshop at each of the available intervals. The function f is a penalty which increases with this distance. H ensures that the desired distance between maintenance checks is not greatly exceeded. The θ_h, $h \in H$ are distances travelled by locomotives at the end of the planning period. The function g is a penalty which increases with the sum of these distances; it ensures that maintenance is not postponed unreasonably to the next period.

Various types of function f can be used. We consider functions of the following family:

$$f_n(\theta_k) = \frac{1}{n} \frac{C}{(5000)^{n+1}} \theta_k^n \qquad (11)$$

These functions have the advantage that they produce a spacing of 5000km between maintenance stops if the sum of maintenance costs per km and penalty costs defined by f are minimized in the absence of network constraints. In fact, if we consider the locomotive being serviced during interval k, the cost per km is C/θ_k and the penalty is $f_n(\theta_k)$. At θ_k = 5000km, we have

$$\frac{d}{d\theta_k}\left(C/\theta_k + f_n(\theta_k)\right)$$

$$= -\frac{C}{\theta_k^2} + \frac{C}{(5000)^{n-1}} \theta_k^{n-1} \Bigg|_{5000\text{km}}$$

$$= -\frac{C}{(5000)^2} + \frac{C}{(5000)^2} = 0$$

The higher the value of n, the faster the penalty increases for θ_k > 5000km.

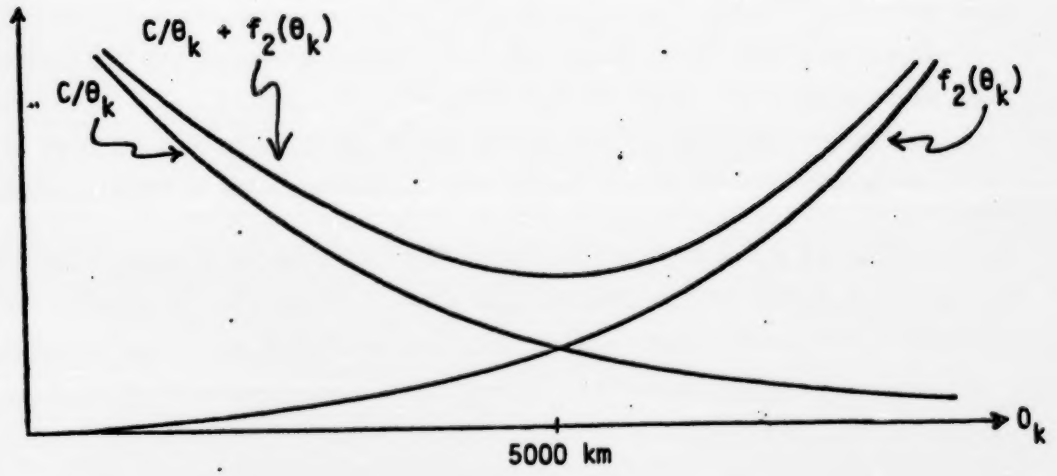


Figure 5: Calibration of penalty f

For the function g, we propose the family

$$g_n(S) = \begin{cases} \frac{1}{A}(S-p)^n & \text{if } S \geq p \\ 0 & \text{otherwise} \end{cases}$$

The parameter A is used to calibrate g in relation to f. An increase of 1000km in θ_h does not have the same importance as an increase of 1000km in θ_k so the value of g must therefore be smaller. The parameter n determines the growth rate of g. Finally, the parameter p represents the value above which the penalty begins to grow. Each week, p could be estimated as follows:

p = total km at beginning of the week + total km covered during the week - (average km between maintenance stops) * (expected number of maintenance jobs to be done during the planning period)

The difference between the second and third terms of this estimate of p should be zero on average over several weeks. However, it could be positive if less maintenance work than usual is to be done during the week being considered because of holidays in the workshop or peak periods requiring the maximum possible circulation of equipment. On the other hand, the difference

could be negative if there is a desire to do more maintenance work than usual in preparation for upcoming holidays in the workshop or peak periods for equipment use. There would be some advantage in making several runs of the program each week varying p and other parameters to produce a number of assignments. Users would then select one according to their preferences.

An objective function of the form $F_1(Y) + F_2(\theta(X))$ must be minimized subject to network constraints. Note that relations (5), (6) and (7) no longer appear as constraints but are used in the calculation of F_2. The objective function is non-linear and piecewise differentiable. This problem can be dealt with using a primal method, beginning with a feasible solution and modifying the solution using the marginal costs so as to reduce $F_1 + F_2$. The optimality of the solution obtained will depend on the convexity of F_2.

4. CONCLUSION

This analysis demonstrates the feasibility of using mathematical methods to deal with the locomotive assignment problem with maintenance constraints. It shows that feasible solutions can be economically derived. The degree of optimality achieved will depend on problem size and complexity and on the level of computation costs considered acceptable.

REFERENCES

(1) Desrosiers, J., Pelletier P. Soumis, F., "Routes avec contraintes d'horaires", HEC #81-20, Ecole des Hautes Etudes Commerciales, Montréal, 25 p. 1981.

(2) Desrosiers J., Pelletier, P., Soumis, F. "Plus court chemin avec contraintes d'horaires", HEC #81-21, Ecole des Hautes Etudes Commerciales Montréal, 20p., 1981.